

How bilinguals listen in noise: linguistic and non-linguistic factors*

JENNIFER KRIZMAN

*Auditory Neuroscience Laboratory
Department of Communication Sciences and Disorders,
Northwestern University*

ANN R. BRADLOW

Department of Linguistics, Northwestern University

SILVIA SIU-YIN LAM

*Auditory Neuroscience Laboratory
Department of Communication Sciences and Disorders,
Northwestern University*

NINA KRAUS

*Auditory Neuroscience Laboratory
(www.brainvolts.northwestern.edu)
Department of Communication Sciences, Northwestern
University
Institute for Neuroscience, Northwestern University
Department of Neurobiology and Physiology, Northwestern
University
Department of Otolaryngology, Northwestern University*

(Received: August 18, 2015; final revision received: March 05, 2016; accepted: March 7, 2016)

Bilinguals are known to perform worse than monolinguals on speech-in-noise tests. However, the mechanisms underlying this difference are unclear. By varying the amount of linguistic information available in the target stimulus across five auditory-perception-in-noise tasks, we tested if differences in language-independent (sensory/cognitive) or language-dependent (extracting linguistic meaning) processing could account for this disadvantage. We hypothesized that language-dependent processing differences underlie the bilingual disadvantage and predicted that it would manifest on perception-in-noise tasks that use linguistic stimuli. We found that performance differences between bilinguals and monolinguals varied with the linguistic processing demands of each task: early, high-proficiency, Spanish–English bilingual adolescents performed worse than English monolingual adolescents when perceiving sentences, similarly when perceiving words, and better when perceiving tones in noise. This pattern suggests that bottlenecks in language-dependent processing underlie the bilingual disadvantage while language-independent perception-in-noise processes are enhanced.

Keywords: listening in noise, adolescence, sensory processing, cognitive processing

Introduction

Typical, conversational speech is a continuous, rapidly unfolding signal that contains variability and ambiguity throughout the duration of the utterance (McClelland & Elman, 1986). Perception of this complex signal is a multi-step process that includes accurate sensory processing of the signal and matching utterances contained within

it to their correct phonological, lexical, and semantic representations (Lecumberri, Cooke & Cutler, 2011; McClelland & Elman, 1986; Norris & McQueen, 2008). Selecting the correct representations from the wealth of potential targets requires that lateral and top-down feedback mediate competition between the eventual linguistic targets and their phonological, lexical, and semantic competitors (Anderson & Kraus, 2010; Norris & McQueen, 2008; Shook & Marian, 2013). Though this speech recognition process occurs with relative ease in quiet listening conditions, it becomes more challenging when the utterance occurs in acoustically-complex environments, such as when speech is presented in noise (e.g., see Shi, 2010). Noise can degrade or mask the rapidly presented, ambiguous cues (Bronkhorst, 2000; Lecumberri & Cooke, 2006). As a result of this degradation, the cues may now match a larger number of

*The authors thank the members of the Auditory Neuroscience Laboratory for their help in data collection and distillation. The authors also thank Jessica Slater, Trent Nicol, and Travis White-Schwoch for their helpful comments on an earlier version of the manuscript and Christine Junhui Liu for her help in copyediting the manuscript. This research is funded by NSF SMA1015614, NIH DC009399, NIH 5F31DC014221, Cognitive Science Program, Northwestern University, and the Knowles Hearing Center, Northwestern University.

Address for Correspondence:

Nina Kraus, Ph.D., Northwestern University, 2240 Campus Drive, Evanston IL, 60208, USA
nkraus@northwestern.edu

potential phonemic or semantic targets than when speech is spoken in quiet. This lack of signal clarity and the need to resolve heightened ambiguity of the rapidly-unfolding utterance makes speech-in-noise perception a complex, challenging process.

Noise makes speech perception challenging for everyone; however, it may result in greater adversity for bilinguals than monolinguals. Much of the evidence for a bilingual disadvantage comes from studies assessing sequential language learners (i.e., late bilinguals) on speech-in-noise perception abilities in their second, non-native language (L2) (e.g., Bradlow & Alexander, 2007; Mayo, Florentine & Buus, 1997; Rogers, Lister, Febo, Besing & Abrams, 2006). In these studies, bilinguals require greater signal resolution, for example through a larger signal-to-noise ratio (Shi, 2010, Mayo et al., 1997) or an increase in both the clarity and predictability of the speech signal (Bradlow & Alexander, 2007) than monolinguals. The level of clarity necessary for perception tracks with age of acquisition and proficiency in the target language such that later age of acquisition or lower proficiency require greater levels of signal clarity (Brouwer, Van Engen, Calandruccio & Bradlow, 2012; Shi, 2010, 2012; Van Engen & Bradlow, 2007). However, these factors alone cannot fully account for the bilingual speech-in-noise disadvantage. The influence of proficiency and acquisition age on L2 speech perception often manifests as deficits in both quiet and noise (Shi, 2010); and yet, even early, near-native proficiency bilinguals who perform equivalently to monolinguals on perceiving speech in quiet show a larger performance drop than their monolingual peers when perceiving speech in noise (Rogers et al., 2006; Shi, 2010). Thus, it is likely that differences in how bilinguals and monolinguals process speech in noise also contribute to the observed performance differences.

The speech-in-noise disadvantage ultimately manifests as poorer utilization of the contextual cues present in the degraded utterance (e.g., using other words in the sentence to identify an unknown word; Cooke, Lecumberri & Barker, 2008; Lecumberri et al., 2011). However, because speech-in-noise perception is a complex, multi-step process, it has been difficult to pinpoint why bilinguals are poorer in utilizing contextual cues. This disadvantage could stem from a number of different processing sources that can be broadly divided into two categories: language-independent processes and language-dependent processes. Language-independent processes are necessary for speech-in-noise perception but are independent of language knowledge, including sensory processing of the signal, separation of the target stream from competing streams, executive control of attention, and working memory. Though the sensory and cognitive processes involved in speech-in-noise perception are tuned by language experience (Krishnan,

Xu, Gandour & Cariani, 2005; Kroll & Bialystok, 2013) or may be called upon during linguistic processing, ultimately they are precursors to language comprehension. On the other hand, language-dependent processes rely on language knowledge, including activation of potential linguistic targets, selection of correct phonological, lexical, or semantic targets from their competitors, and facility in choosing the correct lexical/semantic targets. (Bradlow & Alexander, 2007; Bradlow & Bent, 2002; Lecumberri et al., 2011). Identifying the unique contributions of language-independent and language-dependent processes can provide mechanistic insights into why monolinguals and bilinguals perform differently on tests of speech-in-noise perception.

A bilingual is not the sum of two monolinguals (Grosjean, 1989) and, so, knowing and communicating across two languages necessitates different speech-processing strategies (Shook & Marian, 2013). These differences in speech processing result in certain bilingual advantages for language-independent sensory processes, such as enhanced neural encoding of auditory stimuli (Krizman, Marian, Shook, Skoe & Kraus, 2012; Krizman, Skoe, Marian & Kraus, 2014; Krizman, Slater, Skoe, Marian & Kraus, 2015) and language-independent cognitive processes, including filtering out irrelevant sensory information (i.e., inhibitory control; Bialystok, 2009; Bialystok, 2011; Krizman et al., 2012). Though bilingual experience seems to enhance some language-independent processes, we hypothesize that juggling two languages in one mind creates greater language-dependent processing demands for bilinguals, especially those processes important for lexical access of words in a sentence. If so, then bilinguals should not perform as well as monolinguals when extracting linguistic meaning from a degraded utterance. Conversely, if this degraded stimulus is non-linguistic, bilinguals may outperform monolinguals, given the bilingual advantages in cognitive and sensory processing (Bialystok, 2011; Carlson & Meltzoff, 2008; Krizman et al., 2012; Krizman et al., 2014; Kroll & Bialystok, 2013). Alternatively, if the speech-in-noise disadvantage arises from language-independent processes, monolinguals should outperform bilinguals on all tests that measure perception of a degraded auditory stimulus, even if the target is non-linguistic.

To test whether language-independent or language-dependent processing differences underlie speech-in-noise differences, we compared early (i.e., both languages acquired by the age of 5), high-proficiency Spanish–English bilingual and English monolingual adolescents on English-language tests of sentence-in-noise, word-in-noise, and tone-in-noise perception. By testing early bilinguals that were matched to monolinguals on English proficiency, the effects of proficiency and acquisition age on speech-in-noise performance could be minimized. Moreover, because only early bilinguals, who learned

both languages before the close of the presumed sensitive period (Werker & Tees, 2005), were tested any differences in performance could not be attributed to bilinguals learning a second language outside the sensitive period (Florentine, 1985) and this reduced the likelihood of differences in performance being attributable to native language biasing non-native phoneme perception (reviewed in Lecumberri et al., 2011). Adolescents were chosen because the sensory, cognitive, and linguistic systems that support speech in noise perception (Krizman, Tierney, Fitzroy, Skoe, Amar & Kraus, 2015a; Paus, 2005; Skoe, Krizman, Anderson & Kraus, 2015), and consequently speech-in-noise perception, (e.g., Talarico, Abdilla, Aliferis, Balazic, Giaprakis, Stefanakis, Foender, Grayden & Paolini, 2006) continue to mature during this age. Capturing a diverse range of performance may maximize differences between bilinguals and monolinguals on these perception-in-noise tasks that may not be apparent in young adults. A variety of perception-in-noise measures were chosen because they have a different balance of language-dependent and language-independent processing demands as a consequence of the different amount of contextual cues present in each task's target stimulus. While all tests require language-independent processing, tone perception tests require no language-dependent processing, the word perception test requires some language-dependent processing, and sentence perception tests require the greatest amount of language-dependent processing.

Methods

Participants

Participants were fifty-six freshmen (14.6 ± 0.42 years of age) recruited from three public high schools in Chicago. The Northwestern University Institutional Review Board approved all procedures, and informed written assent and consent was given by the children and their parent/guardian, respectively.

Participants were divided into two groups based on their language experience as measured by the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld & Kaushanskaya, 2007): English monolinguals ($n = 31$; 52% female, 52% low maternal education (i.e., \leq high-school graduate, used as a proxy for socioeconomic status, D'Angiulli, Herdman, Stapells & Hertzman, 2008) and Spanish-English bilinguals ($n=25$; 40% female; 56% low maternal education). Inclusionary criteria were: high English proficiency (≥ 7 out of 10 on English speaking and understanding proficiency, LEAP-Q), low Spanish proficiency for English monolinguals (≤ 3 out of 10 on Spanish speaking and understanding proficiency, LEAP-Q), high Spanish proficiency for bilinguals (≥ 7 out of 10 on Spanish

speaking and understanding proficiency, LEAP-Q); early acquisition of Spanish and English (≤ 5 years old) for bilinguals, air conduction thresholds of < 20 dB hearing level (HL) per octave for octaves from 125–8000 Hz (ANSI, 2009), and no diagnosis of a reading or language disorder.

The two groups did not differ in age (monolinguals: 14.5 ± 0.37 years, bilinguals: 14.6 ± 0.39 years; $F(1,54) = 0.681$, $p = .413$), sex (Kruskal-Wallis $X^2 = 0.737$, $p = .391$), maternal education level (Kruskal-Wallis $X^2 = 1.067$, $p = .302$), IQ (monolinguals: 98.84 ± 10.9 ; bilinguals: 99.2 ± 7.6 ; $F(1,54) = 0.125$, $p = .901$, Wechsler Abbreviated Scale of Intelligence, WASI, Wechsler, 1999), or English proficiency ($F(1,54) = 2.944$, $p = .092$), as determined from the LEAP-Q. The groups were also matched on their English comprehension abilities ($F(1, 54) = 0.320$, $p = .574$; LEAP-Q) but they did differ on daily English/Spanish exposure ($F(1,54) = 222.556$, $p < .0005$) and Spanish proficiency ($F(1,54) = 765.001$, $p < .0005$). Eight of the bilinguals learned English and Spanish simultaneously, 3 were exposed to English first, and the remaining 14 were exposed to Spanish first. For these 14, they learned English on average 2.7 years after learning Spanish. All bilingual participants reported acquiring Spanish at a younger age than they acquired English ($t(24) = 2.477$, $p = .021$), greater proficiency in English than Spanish ($t(24) = 4.299$, $p < .0005$), and greater daily exposure to English than Spanish ($t(24) = 3.239$, $p = .003$). Means and standard deviations of these LEAP-Q measures for each group are summarized in Table 1.

Psychophysical Testing

Participants' listening-in-noise abilities were assessed with two tests of sentence-in-noise perception, one test of word-in-noise perception and two tests of tone-in-noise perception. Two measures of sentence-in-noise and tone-in-noise perception were used to provide converging evidence of any observed differences in performance between monolinguals and bilinguals.

Quick Speech in Noise

The Quick Speech-In-Noise test (QuickSIN, Etymotic Research, Elk Grove, IL; Killion, Niquette, Revit & Skinner, 2001) measured perception of target sentences (female talker) spoken amid four-talker babble (three female and one male). English sentences and babble were presented diotically through insert earphones (ER-2, Etymotic Research, Elk Grove Village, IL). Of the 20-list corpus, three lists were presented to the participant at a fixed level of 70 dB HL. Each list contained six sentences that were syntactically and semantically correct (e.g., 'A white silk jacket goes with any shoes', Wilson, McArdle & Smith, 2007). The signal-to-noise ratio (SNR)

Table 1. *Language Experience Descriptive Statistics*

		English	Spanish
Monolingual n = 31, 52% Female	Native Language	100%	None
	Age of Acquisition	1.5 ± 1.1 years	n/a
	Current Daily Exposure	97.2 ± 5.6 %	2.8 ± 5.6 %
	Proficiency	9.6 ± 0.7	0.6 ± 1.1
Bilingual n = 25, 40% Female	Native Language	32%	68%
	Age of Acquisition	2.8 ± 1.9 years	1.7 ± 1.6 years
	Current Daily Exposure	58.4 ± 13 %	41.6 ± 13 %
	Proficiency	9.2 ± 0.8	8.3 ± 0.9

Table 1. Summary of adolescent monolingual and bilingual language experience as reported on the LEAP-Q.

decreased in 5 dB SNR steps for each sentence in the list with the first sentence presented at 25 dB SNR and the sixth sentence presented at 0 dB SNR. Participants were instructed to repeat back each sentence, and each sentence was scored based on the participant's ability to correctly recall five target words (e.g., "white", "silk", "jacket", "any", "shoes"). The total number of key words correctly recalled in the list (30 in total) was subtracted from 25.5 to give an SNR loss score for each list. The SNR loss score for each list was averaged together to obtain a final score. A lower score indicates better performance on this task.

Hearing in Noise Test

The Hearing In Noise Test (HINT; Biologic Systems Corp., Mundelein, IL; Nilsson, Soli & Sullivan, 1994) was an adaptive test of speech recognition in which the participant repeated short semantically and syntactically simple English sentences (e.g., "A boy fell from the window") presented in speech-shaped background noise that matched the spectra of the test sentences. The Bamford-Kowal-Bench (Bench, Kowal & Bamford, 1979) sentences (12 lists of 20 sentences), spoken by a male, were presented in free field via a loud speaker positioned at 0° azimuth relative to the participant (i.e., directly in front of the participant). Participants were again asked to repeat out loud the target sentence. In this test, the noise presentation level was fixed at 65 dB SPL and the intensity level of the target sentence was increased or decreased to adjust the difficulty level until the threshold SNR was reached. Threshold SNR was defined as the dB SNR difference between the speech dB level and noise dB level that resulted in 50% correct sentence repetition. Again, a lower score indicated better performance on this task. This sentence-in-noise task differed from the QuickSIN in sentence type (simple versus complex), the type of noise (speech-shaped versus babble), type of presentation (speakers versus earphones), signal-to-noise levels (adaptive to threshold versus fixed), and sex of talker (male versus female). By using these different sentence-in-noise tests we could demonstrate whether any

group differences observed were specific to the method of administration or generalizable across tests.

Words in Noise

The Words-in-Noise test (WIN) was a non-adaptive measure of single-word perception, in which participants were asked to repeat a word (e.g., "dog") that was embedded in a carrier phrase (i.e., "Say the word..."). This measure was used to assess perception of single English words in English background babble. The target utterance were spoken by a female voice and masked by four-talker babble (three female and one male) presented at 55 dB HL in soundfield. Thirty-five words were presented starting at a 24 dB SNR and decreasing by 4 dB every five words until 0 dB SNR was reached. The final SNR score was based on the number of correctly-repeated words, corresponding to a threshold performance. As such, lower threshold scores indicated better performance.

Tones in noise

Backward masking and simultaneous masking were used to assess how well participants could perceive auditory stimuli in a task that has no linguistic processing demands. These tests were chosen because they have been shown to activate sensory regions important for speech perception (van Dijk & Backes, 2003). Moreover, the ability to perceive a tone embedded in noise, especially when assessed via backward masking, is known to engage cognitive mechanisms important for segregating rapidly-presented sounds (Hartley, Hill & Moore, 2003; Hartley & Moore, 2002; Hartley, Wright, Hogan & Moore, 2000; Strait, Kraus, Parbery-Clark & Ashley, 2010; Tallal, Miller & Fitch, 1993; Wright, Lombardino, King, Puranik, Leonard & Merzenich, 1997) and corresponds to activation of the prefrontal cortex and anterior cingulate cortex during the task (van Dijk & Backes, 2003), two regions that underlie inhibitory control enhancements found in bilinguals (Abutalebi, Della Rosa, Green, Hernandez, Scifo, Keim, Cappa & Costa., 2011; Kroll & Bialystok, 2013).

These masking tests were assessed by the IHR Multicentre Battery for Auditory Processing (IMAP, developed by the Medical Research Council Institute of Hearing Research, Nottingham, UK; Barry, Ferguson & Moore, 2010) program using a laptop computer placed 60 cm from the participant. Auditory stimuli were presented diotically through Sennheiser HD 25–1 headphones and were accompanied by animated visual stimuli displayed on the laptop screen. A three-button response box was placed in front of the participant and was used to indicate the participant's response on each trial.

For the backward masking test, the participant watched the computer screen while listening to three sequentially-played 'noise sounds', which were composed of a 600 to 1400 Hz bandpass noise (1000 Hz center frequency) that was 300 ms in duration and had a fixed spectrum level of 30 dB. Immediately preceding one of these 'noise sounds', a 20 ms 1000 Hz target tone occurred (i.e., the target's offset and the noise's onset occurred concurrently). On the first trial, the 1000-Hz targets were presented at 90 dB SPL and on subsequent trials, the targets decreased or increased in intensity level via a 3 down, 1 up adaptive staircase model to determine the participants' minimum detection threshold (in dB, see Amitay, Irwin, Hawkey, Cowan & Moore, 2006 for more information on the staircase model). On each trial, participants pressed the button on a 3-button response box that corresponded to the 'noise sound' that contained the target tone (as opposed to noise only).

For the simultaneous masking task, the target tone was a 20 ms, 1000 Hz tone that occurred 200 ms following the onset of a 300 ms noise (the same masking noise used for the backward masking task). The initial target was presented at 95 dB SPL and descended using the same 3-alternative forced-choice paradigm and staircase threshold detection procedure that was applied for the backward masking task. Two blocks of 20 trials each were run for both the backward masking and simultaneous masking tasks, and on both, lower scores indicate better performance. Administration order of backward and simultaneous masking was randomized by the IMAP system. The speech-in-noise tests were always administered in the following order: QuickSIN, HINT, WIN. Practice stimuli were administered at the start of each test to make sure the participant understood the task directions. Administration order of tone-in-noise vs. speech-in-noise tasks was randomized across participants.

Statistical analyses

Each of the perceptual tasks yields a non-standardized threshold score. Since we were interested in comparing performance among these different tests and determining how language experience influenced the relative

performance on these measure, the raw threshold performance values across all subjects in both groups for each test were transformed to z-scores normalized around 0. Comparisons were made between monolinguals and bilinguals on perception-in-noise performance with a 2 (group, monolingual v. bilingual) x 5 (test, QuickSIN, HINT, WIN, backward masking, simultaneous masking) repeated-measures analysis of variance (RMANOVA). Following the RMANOVA analyses, post-hoc t-tests were run comparing monolingual and bilingual performance on the individual perceptual tasks.

For the RMANOVA analyses, corrected p-values are reported in cases where sphericity could not be assumed, as determined by Mauchly's Test of Sphericity, and t-tests report corrected p-values when Levene's Test for Equality of Variances determined unequal variance between the two groups. Cohen's *d* effect sizes are reported for group comparisons where, consistent with accepted standards, a small effect size is ≤ 0.2 , a medium effect size is between 0.2 and 0.5, and a large effect size is ≥ 0.5 .

Results

Overall, monolingual participants outperformed bilinguals on tests of speech perception in noise but bilinguals' tone detection in noise abilities were enhanced relative to monolinguals (Figure 1), as evidenced by a language group by test interaction ($F(4, 51) = 5.845, p = .001$). There was no main effect of language group ($F(1, 54) = 1.335, p = .253$) or test ($F(4, 216) = 0.092, p = .985$). However, because we normalized the performance across the five tests through a z-transformation, a null effect of test was expected.

Post-hoc t-tests showed poorer performance for the bilinguals than monolinguals when perceiving sentences, as measured by both QuickSIN ($t(54) = 2.806, p = .007, d = 0.74$, monolinguals: 1.87 ± 1.5 dB SNR loss, bilinguals: 3.17 ± 1.99 dB SNR loss) and HINT ($t(54) = 3.004, p = .004, d = 0.79$, monolinguals: -0.97 ± 0.82 dB SNR, bilinguals: -0.19 ± 1.12 dB SNR) tests. The groups did not differ on perception of words in noise, WIN ($t(54) = 1.866, p = .087, d = 0.48$), though bilinguals tended to perform more poorly on this measure (monolinguals: 5.51 ± 1.01 dB SNR, bilinguals: 6.22 ± 1.83 dB SNR).

Contrary to poorer performance for bilinguals on the speech-perception-in-noise tasks, bilinguals outperformed monolinguals when perceiving tones in noise. This was evident for both the backward masking task ($t(54) = 2.149, p = .036, d = 0.59$, monolinguals: 53.51 ± 16.67 dB SNR, bilinguals: 44.89 ± 12.4 dB SNR) and the simultaneous masking task ($t(54) = 2.183, p = .033, d = 0.59$, monolinguals: 67.69 ± 5.85 dB SNR, bilinguals: 64.3 ± 5.67 dB SNR). In Figure 1, z-scores are plotted for bilinguals and monolinguals to

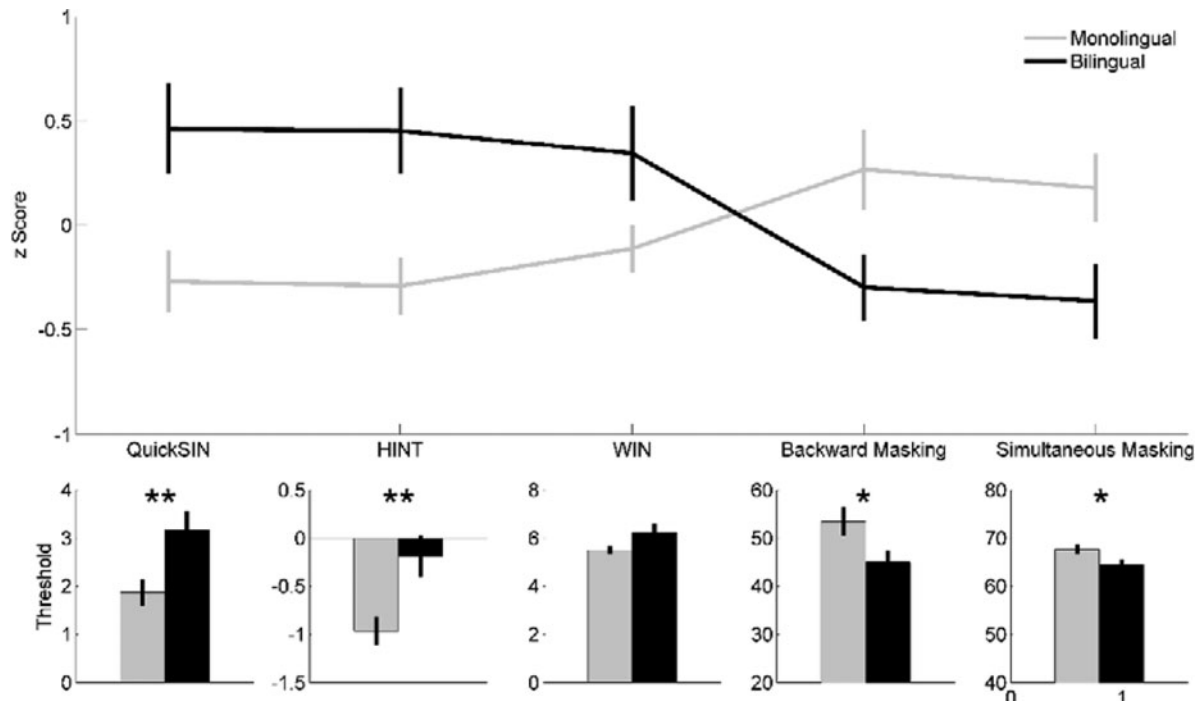


Figure 1. Comparisons of monolinguals (gray) and bilinguals (black) on perception-in-noise tasks. The top plot shows the z-normalized threshold performance for the two groups and the bottom graphs illustrate the raw threshold means (± 1 standard error). Lower numbers indicate better performance for all measures. Monolinguals performed better on the two sentence-in-noise tasks, the two groups performed equivalently on the perception of words, and bilinguals performed better on the two tone-in-noise detection tests.

compare performance across tests. Below the interaction plot, average raw threshold performance for each group is plotted for the different tests.

Discussion

We tested monolingual and early high proficiency bilingual adolescents on sentence-in-noise, word-in-noise, and tone-in-noise perception. We found that differences in performance between bilinguals and monolinguals varied with the amount of linguistic information available in the stimulus. Specifically, we found a monolingual advantage for perceiving sentences in noise and a bilingual advantage when the degraded auditory target is non-linguistic. Our findings of a bilingual speech-in-noise disadvantage in adolescents are consistent with previous studies in young adults (Bidelman & Dexter, 2015; Lecumberri et al., 2011; Mayo et al., 1997; Rogers et al., 2006; Shi, 2010; Shi, 2009). By using non-speech stimuli, the current study also demonstrates that the bilingual perception in noise disadvantage is specific to linguistic stimuli, suggesting that language-dependent processes, but not language-independent processes underlie differences between bilinguals and monolinguals for perceiving speech in noise.

Bilinguals and monolinguals both experience the same signal-driven sources of difficulty when perceiving speech in noise, including degradation of acoustic cues, activation of a greater number of linguistic competitors, and rapidly unfolding speech imposing time constraints on speech-in-noise processing. However, having two languages in one mind requires unique, additional processing for bilinguals relative to monolinguals. While monolinguals only need to resolve competition that arises from within-language competitors (e.g., hearing /k/-/æ/-/n/ can activate both 'candy' and 'candle'), bilinguals experience both within-language and between-language competition. For example, a Spanish–English bilingual hearing /k/-/æ/-/n/ would not only activate 'candy', but Spanish phonological competitors such as 'cántaro' or candy's translation equivalent 'dulce' (Marian & Spivey, 2003a, 2003b; Marian, Spivey & Hirsch, 2003; Shook & Marian, 2013). Moreover, because a phoneme in one language may map with more than one phoneme in the bilingual's other language (e.g., for native Japanese speakers of English, the English phonemes /r/ and /l/; Miyawaki, Jenkins, Strange, Liberman, Verbrugge & Fujimura, 1975), the acoustic signal activates the matching phoneme, as well as words that contain similar-sounding phonemes within either language. Although these overlaps may work asymmetrically across a bilingual's two languages (Cutler,

Weber & Otake, 2006; Pallier, Colomé & Sebastián-Gallés, 2001), they have the potential to cause a cascade of activation for a bilingual that a monolingual would never experience (Lecumberri et al., 2011; Shook & Marian, 2013). Because the presence of two languages in one mind leads to greater language activation, which may be exacerbated when the speech is spoken in noise, bilinguals experience greater linguistic competition and so must devote greater neural resources than monolinguals to competition resolution (Bidelman & Dexter, 2015; Lecumberri et al., 2011; Mattys, Carroll, Li & Chan, 2010).

In addition to contending with a greater amount of competition, bilinguals have not had the same amount of experience with the target language as monolinguals. Therefore, bilinguals do not have the same exposure to the probabilities of co-occurrence that exist for given words, phrases, or syntax within that language (Gollan, Montoya, Cera & Sandoval, 2008; Mack, 1986; Merriman & Kutlesic, 1993; Sorace, 1993). This is important because language-dependent processing is tuned in an experience-dependent manner, such that, when perceiving words in an utterance, knowledge of their prior frequency of occurrence or probability of co-occurrence can be used to identify the correct linguistic targets (Elliott, 1979; Lecumberri et al., 2011; Norris & McQueen, 2008; Talarico et al., 2006). Less experience with the target language appears to result in poorer lexical access and subsequently poorer use of lexical/semantic information during speech perception for bilinguals relative to monolinguals (Mattys et al., 2010; Shook, Goldrick, Engstler & Marian, 2014) and our observation of poorer performance on sentence-in-noise perception is consistent with this difference. Studies investigating relations between speech-in-noise perception and age of acquisition may suggest that this influence is greater in those who have acquired the target language later in life (e.g., Mayo et al., 1997), but these influences were likely at play even in the early age of acquisition participants in the present study.

While many prior studies have found a speech-in-noise disadvantage for sequential, late-learning bilinguals in their second language (e.g., Shi, 2010; Brouwer et al., 2012; Mayo et al., 1997), we add to a growing literature showing (e.g., Rogers et al., 2006) that this difference in speech-in-noise perception is evident in bilinguals who are highly proficient, early learners of the target language. In addition to being early high-proficiency bilinguals, the bilingual group was a mix of Spanish-native, English-native and simultaneous learners. Therefore, we suggest that these effects are the result of bilingualism and not specific to being tested in a non-native (i.e., second) language. Alternatively, since only Spanish–English bilinguals were tested, it is possible that exposure to Spanish may have driven these differences, though we suggest that this is unlikely given that speech-

in-noise differences have been found in bilingual speakers of other languages (e.g., Shi, 2010) and there are no obvious Spanish features that would lead to a tone-in-noise advantage. While these findings suggest that early high-proficiency adolescent bilinguals perform more poorly on tests of speech-in-noise perception and better on tests of tone-in-noise perception relative to their English monolingual peers, future research comparing bilinguals to monolinguals of both languages and studying bilinguals who speak languages different from those spoken by the bilinguals studied here can confirm that the observed differences are an effect of bilingualism and that the speech-in-noise disadvantage is apparent in both of a bilingual's languages rather than a consequence of being tested in a non-native language.

Linguistic-processing bottlenecks may impede bilinguals' perception of speech in noise. However, our finding that bilinguals outperform monolinguals on perception of tones in noise provides evidence for a bilingual advantage in processing degraded non-linguistic stimuli. These results suggest that the bilingual speech-in-noise disadvantage does not result from deficits in language-independent processing. This pattern of a bilingual speech-in-noise disadvantage and non-speech-in-noise advantage suggests that, relative to monolinguals, language-dependent processes are reduced, while language-independent processes are enhanced. In light of previous demonstrations of bilingual cognitive and sensory processing enhancements (Bialystok, 2011; Carlson & Meltzoff, 2008; Krizman et al., 2012; Krizman et al., 2014; Krizman et al., 2015b; Kroll & Bialystok, 2013; Mattys et al., 2010), our pattern of results, together with previous literature (Bidelman & Dexter, 2015), suggest that bilinguals may attempt to compensate, at least partly, for the linguistic processing bottlenecks by enhancing language-independent processes important for perceiving speech in noise. Alternatively, the greater linguistic activation in bilinguals makes perceiving speech in noise more challenging and so it leads to strengthened non-linguistic processing. While differences in linguistic processing demands were the key manipulation in the current study, this manipulation may have also impacted the cognitive processing load required to perform these tasks (Mattys & Wiget, 2011). By obtaining threshold performance measures on the HINT and tone-in-noise tests, task difficulty effects were minimized. Nevertheless, future research should systematically manipulate cognitive load and linguistic load to assess the separate influence of each on bilingual and monolingual speech-in-noise perception.

Adolescents

We have expanded previous findings of a bilingual disadvantage for speech-in-noise perception in adults to

adolescents (e.g., see Mayo et al., 1997; Shi, 2010; Shi, 2009). Follow-up studies should examine whether the bilingual tone-in-noise advantage extends into adulthood. Exploring how bilinguals and monolinguals perform on these tasks in a stable adult system could elucidate whether these language-independent processes are enhanced in bilinguals or if bilingual adolescents reach maturity faster. Furthermore, observing speech-in-noise differences in a younger population in which speech-in-noise perception skills are still developing suggests that the accrual of different language experiences throughout childhood for monolinguals and bilinguals is sufficient to result in differences between these language groups on speech-in-noise tasks. Continuing this investigation into even younger children could help to identify how much single-language or dual-language experience is necessary for the monolingual speech-in-noise advantage and bilingual tone-in-noise advantage.

Clinical Implications

These study outcomes are clinically relevant, as speech-in-noise perception is a chief complaint for patients seeking clinical services (Carhart & Tillman, 1970). Despite the early high-proficiency Spanish–English bilingual adolescents performing significantly worse than the English monolingual adolescents on tests of speech-in-noise perception, both groups still performed within the clinically-normal range on these tests (Killion et al., 2001; Wilson et al., 2007b). Therefore, though they differ in their performance, it is not the case that knowing and using two languages results in impairments in speech processing for bilinguals. Rather, the monolingual/bilingual performance differences observed on the various perception-in-noise tasks suggest that through differences in lifelong language experience, the language-independent and language-dependent processes are differentially shaped, leading to differences in the strategies that bilinguals and monolinguals employ when understanding speech in noise (see Cutler, Lecumberri & Cooke, 2008 for a related discussion of differences in L1 and L2 listening strategies; Cutler, Weber, Smits & Cooper, 2004).

Conclusions

Differences between monolingual and bilingual adolescents in perceiving degraded auditory stimuli vary with the amount of linguistic information available in the stimulus. From this, we confirm previous findings of a bilingual speech-in-noise disadvantage. However, we also find an advantage for bilinguals when detecting non-linguistic sounds in the presence of noise. While future studies should assess the generalizability of the listening strategies observed in the current study by extending this

work to speakers of other languages and other age groups, as well as test participants in both languages they speak, these results suggest that the source of bilingual speech-in-noise perception disadvantage results from bottlenecks in linguistic processing. Our results further suggest that bilinguals may compensate for the language-processing bottlenecks by enhanced cognitive and sensory processes important for perceiving speech in noise. Together, these results highlight how shaping sensory, cognitive, and linguistic processes through experience can lead to advantages and disadvantages in performing real-world tasks.

References

- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2011). Bilingualism Tunes the Anterior Cingulate Cortex for Conflict Monitoring. *Cerebral Cortex*.
- Amitay, S., Irwin, A., Hawkey, D. J., Cowan, J. A., & Moore, D. R. (2006). A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners. *The Journal of the Acoustical Society of America*, 119(3), 1616–1625.
- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *Journal of the American Academy of Audiology*, 21(9), 575.
- Barry, J. G., Ferguson, M. A., & Moore, D. R. (2010). Making sense of listening: the IMAP test battery. *Journal of visualized experiments: JoVE*(44).
- Bench, J., Kowal, Å., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, 13(3), 108–112.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3–11.
- Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, 65(4), 229–235.
- Bidelman, G. M., & Dexter, L. (2015). Bilinguals at the “cocktail party”: Dissociable neural activity in auditory–linguistic brain regions reveals neurobiological basis for nonnative listeners’ speech-in-noise recognition deficits. *Brain and language*, 143, 32–41.
- Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, 121, 2339.
- Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, 112, 272.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica*, 86(1), 117–128.
- Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language

- familiarity and semantic content. *The Journal of the Acoustical Society of America*, 131, 1449.
- Carhart, R., & Tillman, T. W. (1970). Interaction of competing speech signals with hearing losses. *Archives of Otolaryngology*, 91(3), 273–279.
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(2), 282–298.
- Cooke, M., Lecumberri, M. L. G., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, 123, 414.
- Cutler, A., Lecumberri, M. L. G., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *The Journal of the Acoustical Society of America*, 124(2), 1264–1268.
- Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, 34(2), 269–284.
- Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, 116(6), 3668–3678.
- D'Angiulli, A., Herdman, A., Stapells, D., & Hertzman, C. (2008). Children's event-related potentials of auditory selective attention vary with their socioeconomic status. *Neuropsychology*, 22(3), 293–300.
- Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, 66(3), 651–653.
- Florentine, M. (1985). Speech perception in noise by fluent, non-native listeners. *The Journal of the Acoustical Society of America*, 77(S1), S106–S106.
- Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58(3), 787–814.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15.
- Hartley, D. E., Hill, P. R., & Moore, D. R. (2003). The auditory basis of language impairments: temporal processing versus processing efficiency hypotheses. *Int J Pediatr Otorhinolaryngol*, 67 Suppl 1, 137–142.
- Hartley, D. E., & Moore, D. R. (2002). Auditory processing efficiency deficits in children with developmental language impairments. *J Acoust Soc Am*, 112(6), 2962–2966.
- Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *J Speech Lang Hear Res*, 43(6), 1402–1415.
- Killion, M. C., Niquette, P. A., Revit, L. J., & Skinner, M. W. (2001). Quick SIN and BKB-SIN, two new speech-in-noise tests permitting SNR-50 estimates in 1 to 2 min. *The Journal of the Acoustical Society of America*, 109(5), 2502–2502.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25(1), 161–168.
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences*, 109(20), 7877–7881.
- Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, 128, 34–40.
- Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015a). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience Letters*, 585, 48–53.
- Krizman, J., Tierney, A., Fitzroy, A. B., Skoe, E., Amar, J., & Kraus, N. (2015b). Continued maturation of auditory brainstem function during adolescence: A longitudinal approach. *Clinical Neurophysiology*.
- Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25(5), 497–514.
- Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119, 2445.
- Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2011). Non-native speech perception in adverse conditions: A review. *Speech Commun*.
- Mack, M. (1986). A study of semantic and syntactic processing in monolinguals and fluent early bilinguals. *Journal of Psycholinguistic Research*, 15(6), 463–488.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940.
- Marian, V., & Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, 24(02), 173–193.
- Marian, V., & Spivey, M. (2003b). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, 6(2), 97–115.
- Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, 86(1), 70–82.
- Mattys, S. L., Carroll, L. M., Li, C. K., & Chan, S. L. (2010). Effects of energetic and informational masking on speech segmentation by native and non-native speakers. *Speech Communication*, 52(11), 887–899.
- Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145–160.
- Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing Research*, 40(3), 686.

- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.
- Merriman, W. E., & Kutlesic, V. (1993). Bilingual and monolingual children's use of two lexical acquisition heuristics. *Applied Psycholinguistics*, *14*, 229–229.
- Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340.
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*, *95*(2), 1085–1099.
- Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357.
- Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445–449.
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends Cogn Sci*, *9*(2), 60–68.
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(3), 465.
- Shi, L.-F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language and Hearing Research*, *53*(4), 821.
- Shi, L.-F. (2012). Contribution of Linguistic Variables to Bilingual Listeners' Perception of Degraded English Sentences. *Journal of Speech, Language and Hearing Research*, *55*(1), 219.
- Shi, L. F. (2009). Normal-hearing English-as-a-second-language listeners' recognition of English words in competing signals. *International Journal of Audiology*, *48*(5), 260–270.
- Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2014). Bilinguals Show Weaker Lexical Access During Spoken Sentence Comprehension. *Journal of Psycholinguistic Research*, 1–14.
- Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, *1*(1), 1–21.
- Skoe, E., Krizman, J., Anderson, S., & Kraus, N. (2015). Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex*, *25*(6), 1415–1426.
- Sorace, A. (1993). Unaccusativity and auxiliary choice in non-native grammars of Italian and French: asymmetries and predictable indeterminacy. *Journal of French Language Studies*, *3*(01), 71–93.
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res*, *261*(1-2), 22–29.
- Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., Foenander, K., Grayden, D. B., & Paolini, A. G. (2006). Effect of age and cognition on childhood speech in noise perception abilities. *Audiology & neuro-otology*, *12*(1), 13–19.
- Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, *682*, 27–47.
- van Dijk, P., & Backes, W. H. (2003). Brain activity during auditory backward and simultaneous masking tasks. *Hearing Research*, *181*(1), 8–14.
- Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native-and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, *121*, 519.
- Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3), 233–251.
- Wilson, R. H., Carnell, C. S., & Cleghorn, A. L. (2007a). The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *Journal of the American Academy of Audiology*, *18*(6), 522–529.
- Wilson, R. H., McArdle, R. A., & Smith, S. L. (2007b). An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*, *50*(4), 844–856.
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, *387*(6629), 176–178.